The paper by Kalivitis et al. presents long term measurement of particle size distribution from Finokalia (eastern Mediterranean region). The main focus of the study is on nucleation mode particles and characteristics of new particle formation (NPF) events, including frequency of occurrence as well as particle formation and growth rates. The last part of the paper is dedicated to a simulation case study of NPF with the MALTE-box model.

I recommend the publication of this paper, as it is well written and describes a valuable dataset which allows for the investigation of NPF over 7 years, thus contributing to our understanding of the process. I would however suggest some revisions before final publication of this study. In particular, some of the observations/conclusions reported throughout the manuscript should be slightly balanced. Also, I am not fully convinced by the modelling part in its current form: it is in my view missing a clear presentation of the strategy/sensitivity tests which lead to the final "good simulation", and it would also benefit from a quick discussion on the relevance of the values finally used for some of the key variables (e.g. monoterpenes concentration). Moreover, it is not clear to me how the analysis reported in Section 3.5 of the present paper differs from that of Tzitzikalaki et al. (2017), as I cannot access this source. Detailed comments are listed below.

P3, L13-16 : I would suggest to clearly mention "only when <u>accumulation mode</u> particles were neutral", as with the current form of the sentence it is a bit confusing whether those particles are pre-existing particles or the newly formed ones.

P4, L21: I would suggest to remove "from the early stages of nucleation", since I think those cannot be investigated when measuring particles larger than 9 nm. Such statement would better suit to AIS measurements or to measurements conducted with instruments such as the particle size magnifier (PSM, Vanhanen et al., 2011), which allows for the detection of ~1-1.5 nm particles (charged + neutral).

P5, L9-10: Please refer the reader to Mirme et al., (2007) for AIS measurements. Also, could the authors give more information about the uncertainties reported on L14-15 (calculation method or reference to a paper)?

P5, L22-31: Several short/minor comments about the description of the calculations:

- L22: instead of "particles with diameter D" (should at least be Dp) and since the formation rate is not calculated for different particle diameters, I would clearly mention Dp = 9nm, otherwise one has to wait until Section 3.2 to explicitly get this information (and it would also be more consistent with the description of the terms of Eq. (1));
- L25: "CoagS is the coagulation of particles in this size range" (should be $CoagS_{Dp}$): which (lowest) particle size was used to calculate $CoagS_{Dp}$? I would suggest a more accurate formulation, such as "$CoagS_{Dp}$ is the coagulation sink of XX nm particles on larger particles";
- L27: Please refer the reader to Dal Maso et al. (2005) for the mode fitting method;
- L31: For this first occurrence, instead of "the sulfuric acid sink", I would suggest to rather write something more explicit like "CS is the condensation sink caused by the pre-existing aerosol population and was calculated using the characteristics/properties of sulfuric acid".

P6, L11: "relevant chemical reactions": I would recommend to add few words on the relevance of the reactions, at least mention they are related to sulfuric acid production.

P6, L26-27: what is "free form nucleation"? I would suggest to briefly recall the parameterization which is used in the model and introduce the "nucleation coefficient", later discussed in Section 3.5 (P16, L16 & L24-25).

P6, L29-30: "All these compounds were treated as sulfuric acid and organics": what does this sentence mean? Also, on L27, if ELVOCs are considered please add "20 <u>extremely</u> low-volatility organic compounds", otherwise change to "LVOCs."

P7, L6-7: I am a bit confused with this sentence: only the particle size distributions are used to initialize the model (as reported on P6, L24-25), which then calculates a CS based on the simulated distributions, right? If the purpose of the abovementioned sentence is only to precise that SMPS data were used to calculate the CS, I would strongly recommend to move it to Section 2.1 (P5, L31), as Section 2.2 is dedicated to model description.

P8, L11-13: I am a bit confused with the use of TUV: was the parameterisation used instead of TUV, or implemented in TUV?

P8, L21-27: I am somewhat sceptical about the values which are reported in this paragraph; I think they do not give much information since the shape of the particle size distribution is highly variable with respect to seasons, event vs non-event days, time of the day… I would thus suggest to either provide a more detailed description/comparison of the concentrations in the different modes and their contribution to total concentration, or at least provide quartiles/standard deviation for all reported values (not only for nucleation mode particle concentration).

P8, L31: are the times local or UTC?

P8, L32 - P9, L1: "Such an observation suggests that the nucleation particle number concentration is controlled by NPF episodes". Isn't it what we expect by definition? Which other sources would the authors expect for particles in this size range? This comment also refers to P2 L16-17, P9 L25, P11 L21-22, P15 L22-23. Moreover, concerning the statement P11 L20-22, I am not sure if the linear relation between J_9 and nucleation mode particle concentration (N_{9-25}) can be considered as a strong support for NPF being the main source of nucleation particle, since according to Eq. (1) J_9 calculation includes N_{9-25} in two of the three terms.

P9, L10-23: I think that even if deep investigation of night time events is not in the scope of this paper, slightly more detailed description could be provided. In specific:

- L10-12: Even if similar night time concentrations are observed during all seasons, they seem to result from different processes based on Fig. 2a. Indeed, there is an increase of the concentrations after 18:00 in summer and autumn, which may suggest evening time new particle formation, but during spring and winter the concentrations keep on decreasing until they reach the night time value, suggesting that evening events are not frequent during these seasons. I would thus suggest to balance the sentence from L10-12, and maybe provide frequencies of occurrence of such events for each season, which will also help quantifying "Frequently" (L13).
- L17-18: I would also add that on top of the "local" character of these events, which may partly explain the limited source of condensing vapours (and therefore particle growth), the absence of photochemistry during night time most likely strengthen the lack of vapours needed to sustain particle growth.

P10, L3-18: I wouldn't say that ozone is "<u>the</u> major oxidant in the atmosphere", especially when focussing on daytime NPF events, during which OH is expected to play a significant (major?) role. Also, I don't think that based on the variables included in this factor analysis it is possible to state that NPF is not sensitive to "atmospheric chemical composition"; compounds other than ozone such as for e.g. NOx, SO_2, monoterpenes… would be needed to draw such conclusions.

P11, L3: "the particle survival probability <u>seems</u> to be the highest in winter": the authors have the data needed to actually test their hypothesis and provide a more robust conclusion, and even quantify the variations of the survival probability in different seasons.

P11: While they peak at slightly different times of the year, the maximum of the NPF frequency, particle formation and growth rates are all attributed to enhanced biogenic emissions and/or photochemistry (P10 L27, P11 L16-17 and P11 L25-26, respectively). This hypothesis seems plausible as all maxima are observed during spring/summer, but could the authors comment on the different seasonal variations of the abovementioned variables? In contrast it can be seen from Fig. 1a and 6b that the GR and CS have similar seasonal patterns: is it then realistic to think that CS and the vapours involved in particle growth "share the same origin"?

P11, L27-28: It is true that based on Fig 2a the average duration of NPF in summer seems to be shorter compared to other seasons, but also the maximum of the concentration is lower. Since the CS (and consequently CoagS) is also higher during summer (Fig. 1a), I would think that both the CS (CoagS) and the GR are affecting the variation of nucleation mode particle concentration (should be checked by calculating the survival probability).

P12, L1, L5: I would slightly balance the statements ("notable increase", "clear decreasing") as in my opinion the reported observations are not as obvious as suggested.

P12, L5-33: I am not fully convinced by the conclusion reported on L30, which suggests that decreased SO_2 concentrations related to the economic crisis in Europe may explain observed variations of GR and occurrence of class I NPF events. Main reasons for this are listed below:

- The lack of SO_2 measurement in Finokalia prevents from any direct evaluation of the SO_2 concentration decrease at this site;
- Based on previous studies mentioned in the present work it seems that decreasing SO_2 concentrations can lead to contrasting observations, thus pointing to the fact that robust conclusions cannot be inferred from the analysis of SO_2 alone;
- While the important role of H_2SO_4 in early nucleation stage has been reported in different studies, the need for other species to explain observed GR has also been evidenced, and the present paper itself tends to emphasize the role of organic species in NPF at Finokalia. Indeed, maximum of NPF occurrence, J_9 and GR are all attributed to enhanced biogenic emissions, and best agreement between model simulation and observation is achieved when adjusting monoterpenes concentration in the model. I would thus think that SO_2 driving the observed variations of GR and NPF occurrence is not fully consistent with the aforementioned observations/results.

P13, L12-14: Does this sentence mean that instrument malfunction was affecting measurement of positive ions? If not, it is fine to focus on negative ions only, but I wouldn't justify this choice based on their better ability to represent NPF events. Indeed, it is in my opinion complex to assess which polarity gives the "better representation of NPF events", as the different observations from the two DMAs may instead reflect the signature of the nucleation mechanism.

P13, L17-26: I would have expected the AIS-derived NPF frequencies to be more often higher (or at least equal) than SMPS-derived ones, while the opposite is shown on Fig. 10. Does it mean that the event day illustrated on Fig. 9 is only representative of a rather limited fraction of the events observed in Finokalia, while the majority of them is actually not visible from the AIS smallest diameters? In order to make the most of the FRONT dataset and provide more information on the nature of the events detected in Finokalia during this period, I would suggest to also report for each month (on Fig. 10 for instance) the number of events detected by each instrument and the number of event days they have in common. This will help assessing the fraction of events with very limited growth only visible in AIS data, the fraction of regional events detected by both instruments and that of events only visible in SMPS data.

P14, L6-33: I have several comments/questions regarding model simulations:

- L15: What does "NPF level" mean?
- L22-26: Could the authors briefly summarize the strategy they adopted to finally reach fair agreement between model simulation and observation? For instance, which sensitivity tests

were performed, were parameters other than nucleation coefficient and monoterpenes concentration also tuned?

- L25: are the levels of final simulated monoterpenes concentration realistic, are they for instance in agreement with observations from 2014?
- L27-29: I would slightly balance the conclusions (“well captured”, “in such detail”), as if I agree with the fact that the reported results are very encouraging, one can observe some discrepancies between model and observation (e.g. NPF event from day 243 in not visible in model data);
- L29-31: Do the authors also consider the possibility to test other nucleation mechanisms in future simulations?

References:

Mirme, A., E. Tamm, G. Mordas, M. Vana, J. Uin, S. Mirme, T. Bernotas, L. Laakso, A. Hirsikko, and M. Kulmala: A wide-range multi-channel Air Ion Spectrometer, Boreal Environ. Res., 12, 247–264, 2007.

Vanhanen, J., Mikkilä, J., Lehtipalo, K., Sipilä, M., Manninen, H. E., Siivola, E., Petäjä, T., and Kulmala, M.: Particle size magnifier for Nano-CN detection, Aerosol Sci. Tech., 45, 533–542, 2011.